

DOI:10.19651/j.cnki.emt.2211101

多工况下基于 EEMD-ICNN 的输油管道泄漏识别*

骆正山 刘雨静 王小完

(西安建筑科技大学管理学院 西安 710055)

摘要: 针对多工况下管道泄漏信号预处理繁琐、误报率高的问题,提出了一种集合经验模态分解(EEMD)结合改进卷积神经网络(ICNN)的泄漏识别模型。所用识别方法采用 EEMD 将泄漏信号分解成若干个具有稳态性能的固有模态分量(IMF),通过相关系数划分出噪声主导向量并予以去除实现信号重构;提取重构信号的一系列指标特征作为 ICNN 模型的输入进行特征提取,实现管道多工况分类;ICNN 在每个卷积层和池化层之间加入批量归一化层,以此加快网络训练速度。结果表明:所提模型能够快速准确识别出停泵、调阀、泄漏、正常工况,且在较少训练数据下平均识别准确率可达 98.25%。与未改进的 CNN 和 SVM 分类识别模型相比,该方法有效提高了识别准确率。

关键词: 原油集输管道;泄漏信号;管道噪声;集合经验模态分解;卷积神经网络

中图分类号: TE86 **文献标识码:** A **国家标准学科分类代码:** 440.55

Oil pipeline leakage identification based on EEMD-ICNN under multiple working conditions

Luo Zhengshan Liu Yujing Wang Xiaowan

(School of Management, Xi'an University of Architecture and Technology, Xi'an 710055, China)

Abstract: Aiming at the problems of cumbersome pre-processing and high false alarm rate of pipeline leakage signals under multiple working conditions, an ensemble empirical mode decomposition (EEMD) combined with improved convolutional neural network (ICNN) is proposed as a leakage identification model. The proposed identification method uses EEMD to decompose the leak signal into several intrinsic modal components (IMFs) with steady-state performance, and the noise dominant vectors are divided and removed by correlation coefficients to achieve signal reconstruction. A series of indicator features of the reconstructed signal are extracted as the input of the ICNN mode for feature extraction to achieve the multi-condition classification of pipeline. The batch normalization layer was added by ICNN between each convolutional layer and the pooling layer to accelerate the network training. In accordance with the results, it indicates that the proposed model can quickly and accurately identify pump shutdown, valve adjustment, leakage and normal operating conditions, which can reach 98.25% of the average recognition accuracy under less training data. This technique significantly raises the accuracy of recognition when compared to the unimproved CNN and SVM classification recognition models.

Keywords: crude oil gathering and transportation pipeline; leakage signal; pipeline noise; ensemble empirical mode decomposition; convolutional neural network

0 引言

输油管道在实际运行过程中常受到来自人为破坏、管材腐蚀、自然灾害等多种原因造成泄漏事故频发^[1],不仅造成了环境污染,还会给国家带来巨大的经济损失。管道泄漏检测现场往往存在调阀、停泵等正常操作引起的压力变化干扰识别准确率。因此及时准确地诊断出管道故障对保

障人员、环境和经济安全等具有重要意义^[2]。

为保障输油管网安全运行,降低多工况调节带来的误报警率,国内外学者在信号检测方面进行了大量的研究工作^[3-5]。根据信号检测原理可分为非声学 and 声学检测法^[6],基于声学原理的检测方法具有灵敏度高、误报率低等优势,已有学者将小波变换(wavelet transform, WT)^[7]、经验模态分解(empirical mode decomposition, EMD)^[8]、集合经验

收稿日期:2022-08-17

* 基金项目:国家自然科学基金(41877527)、陕西省社科基金(2018S34)项目资助

模态分解 (ensemble empirical mode decomposition, EEMD)^[9] 等方法用于信号预处理。王晓敏等^[10] 利用小波变换与统计技术融合, 成功提取出了泄漏信号的特征量。魏强等^[11] 提出一种以 EMD 算法为基础的自适应降噪方法, 相较于小波变换而言, 它对泄漏信号更加敏感, 网络识别效果更好, 但 EMD 在分解过程中存在模态混叠、端点效应等缺点。Xie 等^[12] 通过 EEMD 对管道声波信号进行降噪处理, 不仅能够保留更多有效的特征信息, 还克服了传统 EMD 算法模态混叠现象, 提高了预测准确率。1997 年唐秀家等^[13] 首次采用神经网络研究管道泄漏问题, 随着智能管网的发展, 人工智能检测方法逐渐被更多的学者重视^[14-15]。Wu 等^[16] 基于声信号技术构建了多层感知神经网络泄漏识别模型, 提高了定位精度。温江涛等^[17] 提出利用压缩感知结合卷积网络提取特征, 最后经 Softmax 输出泄漏孔径分类结果, 虽然该方法识别准确率较高, 但网络模型结构复杂, 训练效率仍有待优化。张瑞程等^[18] 通过建立卷积神经网络直接作用于管道原始一维数据进行特征自提取, 实现了端到端的精准故障诊断。但直接将原始数据输入到网络模型中会增加网络训练时长, 不能充分发挥卷积神经网络强大的特征自提取优势。

综上, 本文提出一种 EEMD 结合改进卷积神经网络 (improved convolutional neural network, ICNN) 的管道泄漏识别模型。将 EEMD 引入到泄漏信号中去噪声主导向量, 解决原始信号冗余的问题; 通过提取可用于信号表征的特征指标作为 ICNN 的输入来区分管道各运行工况, 降低系统误报率; 本文考虑到模型规模较小, 故在 CNN 原有的卷积层和池化层之间引入批量归一化层 (batch normalize, BN), 避免网络出现过度拟合; 与管道泄漏识别领域其它模型进行对比, 验证了所提方法在多工况识别方面的优越性。

1 方 法

1.1 集合经验模态分解 (EEMD)

EEMD^[19] 是一种为解决 EMD 方法模态混叠现象的改进算法。其分解原理是利用高斯白噪声振幅标准偏差恒定的特性将 n 个互不相关且均值为零的白噪声引入待分解的信号中, 白噪声使泄漏信号分解为多个具有稳态性能分量 (intrinsic mode function, IMF), 而各互不相关的高斯白噪声会在多次平均的过程中抵消, 达到抑制信号模态混叠的效果。最终经 EEMD 分解后可以有效地滤除信号中不必要的分量, 将占主导地位 IMF 分量重新组合生成重构信号。

1.2 卷积神经网络 (CNN)

传统卷积神经网络结构如图 1 所示。

输油管道运行环境复杂且信号之间具有较强的非线性关系, 原始数据中包含大量的噪声无法较好的反映信号变化, 因此提取信号中的有效特征是保证识别准确率的前提。本文在传统 CNN 基础上, 利用卷积核执行卷积计算, 对管

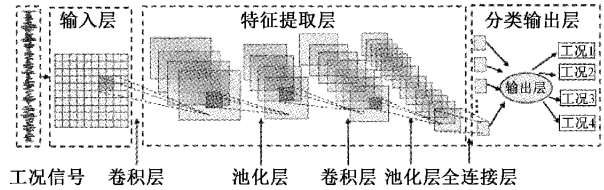


图 1 卷积神经网络基本结构

道信号进行特征提取和稀疏处理; 采用 ReLU 作为激活函数对卷积层输出结果进行非线性表达, 以提高模型的特征学习能力; 引入批量归一层对内部数据进行标准化处理, 以加速网络收敛速度; 最后由全连接层整合管道各工况局部特征, 通过 Softmax 分类器对提取到的特征进行分类输出。

2 实验研究

多工况下基于 EEMD-ICNN 管道泄漏识别模型流程如图 2 所示。

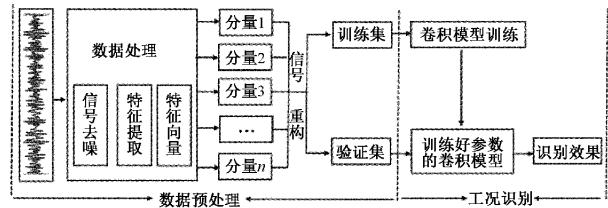


图 2 模型流程

首先实验室搭建模拟某原油集输管道的泄漏场景采集数据, 通过加入幅度系数为 0.1 的白噪声 50 次将管道各工况信号分解为 8 个 IMF; 其次利用相关系数对无效信号进行滤除后重构信号; 最后将训练好的 ICNN 模型作为特征提取工具实现工况分类, 保存训练好的模型用于验证识别效果。

2.1 实验设计

为验证本文所提方法的有效性, 对某原油集输管道的一部分进行泄漏监测, 实验室搭建管道实验系统如图 3 所示。选用钢质管道为研究对象, 管道外径 250 mm, 壁厚 6 mm, 总长 38 m, 正常输油量为 100~300 m³/h。液罐与管道首端之间设置有调节阀, 用于调节介质流速与管压。监测点分别位于管道首末两端, 3 个泄漏孔距首端距离依次为 8、18、27 m。

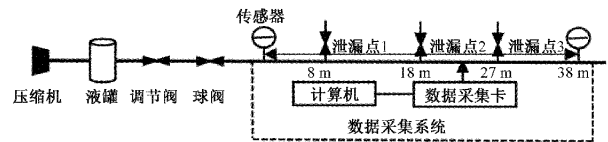


图 3 管道泄漏实验系统

2.2 信号去噪

采集管道信号时, 泄漏信号通过依次打开 3 个不同位置的泄漏孔获取振动数据; 调阀信号通过设置调节阀不同开度获取阀门变化前后的信号数据; 而停泵信号则是利用压缩机和缓冲罐模拟离心泵为实验提供动力, 当失去动力

时获取到的数据变化即为停泵信号;最后管道正常运行状态采用连续采集数据后再分割的方法获取,每类工况各采集100组数据,每组数据采样时间间隔为2min。

实验采集现场往往易受到环境和管内各种噪声的影响,埋地输油管道还会受到地表机械施工等产生的振动信号干扰,所以结合管道实际情况添加高斯白噪声模拟实际

采集信号时的噪声干扰。为了区分泄漏与其他正常工况操作引起的信号波动,对采集到不同周期的振动信号加入幅度系数为0.1的高斯白噪声50次进行集合经验模态分解,以漏点1的数据为例进行EEMD分解,正常、泄漏、调阀、停泵信号依次被分解为8个不同波动尺度的IMF分量如图4所示。

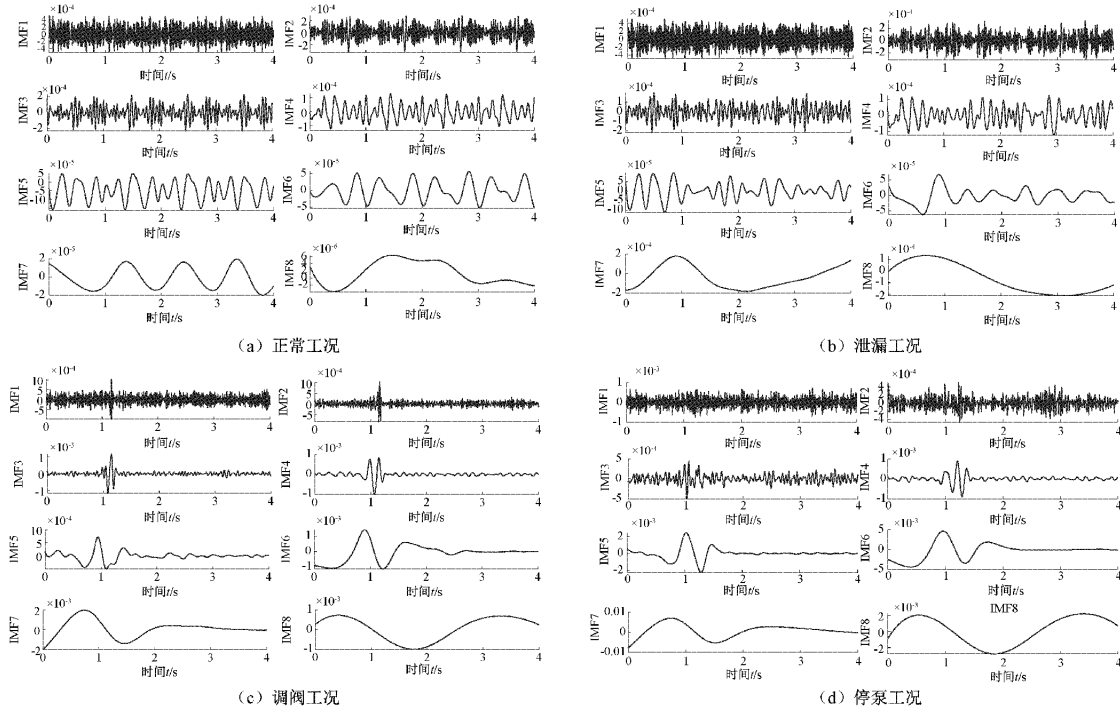


图4 信号EEMD处理结果

由图4可知,信号波动的幅度随着IMF分量阶数的增加而逐渐减小,残差分量和部分低频分量作为特征元素将会干扰识别准确率。为了有效消除管道信号中存在的噪声,根据式(1)计算原始管道信号与各IMF分量的相关系数,记为 ρ_{x, IMF_i} 。

相关系数计算公式如下:

$$\rho_{x, IMF_i} = \frac{\sum_{n=1}^N x(n) IMF_i(n)}{\sqrt{\sum_{n=1}^N x^2(n) \sum_{n=1}^N IMF_i^2(n)}} \quad (1)$$

IMF分量筛选公式如下:

$$\eta = \xi / \delta \quad (2)$$

式中: ξ 为所求取相关系数的最大值, δ 为大于1的比例系数,文中选用 $\delta = 10^{[9]}$ 。当 $\rho_{x, IMF_i} \geq \eta$ 时保留第*i*个向量, $\rho_{x, IMF_i} < \eta$ 时则去除。

相关系数计算结果,如表1所示。

以正常工况为例进行分析,选取相关系数的最大值 ξ 为0.797,根据式(2)求得 η 的值为0.0797,与其余系数对比后,IMF1~IMF6的相关系数值均大于 η 的值,故保留前6个分量进行信号重构,将剩余分量予以去除。其他工况均

采取同样的方法对含噪分量进行滤除后得到重构信号。

表1 各分量相关系数

工况类型	正常	泄漏	调阀	调泵
IMF1	0.797	0.833	0.069	0.011
IMF2	0.493	0.279	0.060	0.096
IMF3	0.314	0.181	0.761	0.198
IMF4	0.295	0.121	0.069	0.315
IMF5	0.271	0.056	0.116	-0.090
IMF6	0.132	0.073	-0.217	-0.229
IMF7	0.035	-0.182	0.312	0.276
IMF8	0.011	-0.056	0.417	0.427

为减少输入数据带宽,降低卷积神经网络的计算量,根据特征参数选取规则^[20]表2给出了6个特征参数,分别为平均幅值、方根幅值、均方根、峭度、峭度因子、裕度因子。从中选取不同测点、不同采集周期的训练样本各200组,测试样本各40组,使用改进卷积神经网络对6种特征参数组成样本集进行特征提取。

2.3 改进卷积神经网络泄漏识别

卷积神经网络模型由特征提取部分和特征映射部分组成,设置合适的参数有利于获取到更多有效的特征。本

表 2 特征指标

特征指标	计算式
平均幅值	$ x = \frac{1}{N} \sum_{i=1}^N x_i $
方根幅值	$X_r = \left(\frac{1}{N} \sum_{i=1}^N \sqrt{ x_i } \right)^2$
均方根	$X_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$
峭度	$\beta = \frac{1}{N} \sum_{i=1}^N x_i^4$
峭度因子	$k = \frac{\beta}{x_{rms}^4}$
裕度因子	$L = \frac{x_p}{x_r}$

表 3 改进卷积神经网络模型参数

参数	取值	参数	取值
卷积层 2	大小[2, 1]、步长[2, 1]	全连接层 1	100
卷积层 3	大小[2, 1]、步长[2, 1]	全连接层 2	4
通道数	128, 64	Softmax 层	—
激活函数	ReLU	迭代次数	200
池化层	大小[2, 2]、步长[1, 1]	训练集比例	80%
学习率	0.001 按 0.5 倍递减	损失函数	Loss

文使用的 ICNN 模型关键参数如表 3 所示。特征提取部分采用 ReLU 作为激活函数(K)。如式(3)所示。

$$k = F_{ReLU}(x) = \max(0, x) \tag{3}$$

式中:(x)为激活函数,x 为输入元素值。

特征映射部分采用 Softmax 分类器输出分类标签,表达式如式(4)所示。

$$\text{Softmax}(Z^aj) = \frac{e^{z^{(j)}}}{\sum_{k=1}^M e^{z^{(k)}}} \tag{4}$$

式中:M 为类别总数,(j)为第 j 神经元输出值。

为避免模型收敛过慢及过拟合现象,批次训练数据将通过随机打乱、随机抽样等方式从样本集中随机抽取进入输入层,通过逐层减少卷积核数目强化。训练集用来对 ICNN 进行训练和参数调优,训练完成后用验证集对 ICNN 的泄漏识别性能进行验证和结果可视化。

3 实验结果与分析

3.1 算法性能分析

实验通过改变管道压力、泄漏孔径大小、泄漏孔距传感器的距离共收集 4 类工况数据 400 组,4 类工况对应标签类别如表 4 所示,从中随机抽取 80%数据作为模型训练集,20%数据作为模型验证集。

将训练样本输入到 ICNN 网络模型中进行训练。模

表 4 样本标签类别

管道工况	训练样本数	测试样本数	标签类别
调阀	80	20	1
调泵	80	20	2
泄漏	80	20	3
正常	80	20	4

型初始学习率为 0.001,每 3 轮训练后学习率按 0.5 倍递减,该操作能使模型在前期训练较快,且后期波动性被明显抑制,从而更加容易接近最优解。使用 Adam 优化器优化模型,交叉熵(Loss)作为损失函数,多次训练后通过图 5 准确率与损失变化曲线来评估模型结构的有效性。

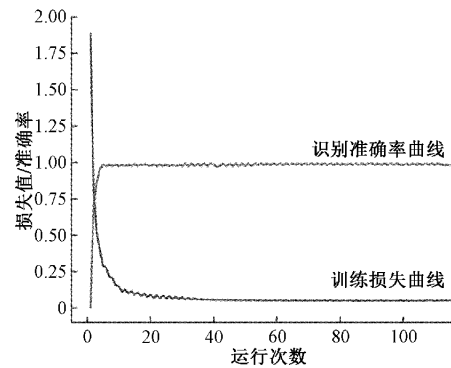


图 5 模型训练收敛曲线

如图 5 所示,随着运行次数的增加,模型的诊断误差在逐渐减小,同时代表模型识别准确率的曲线不断上升。当迭代次数达到 50 次以后,模型训练曲线均开始逐渐平稳,此时模型的损失值为 0.061,识别准确率达到 98.83%。用测试集验证得到识别准确率为 98.25%,说明模型具有较好的识别效果。

研究为了深入了解模型准确率损失在哪些工况上,引入混淆矩阵对各类工况样本的目标物识别进行分析。由图 6 可以看出模型对标签 3、标签 4 运行工况的识别准确率达到 100%,仅有少量的标签 2 在识别时发生了故障,其中原为调泵的一个样本被错误识别为调阀,其余泄漏场景均预测正确。模型整体识别准确率达 98.8%。

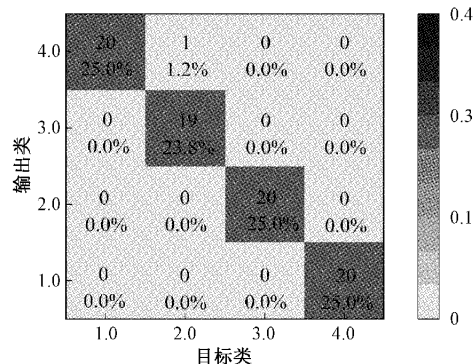


图 6 混淆矩阵

为了进一步研究模型对于管道信号的自适应学习能力,采用t分布(t-distributed stochastic neighbor embedding, t-SNE)算法对网络各层级特征进行可视化分析^[21],如图7所示。输入层各工况样本分布杂乱无法进行区分,经第一

卷积层提取特征后,多数样本开始在区域内聚集,第2层特征样本聚集度明显增高,逐层卷积完成后样本特征高度区分,实验结果表明,EEMD-ICNN模型能够从原始信号中提取有用特征,实现准确的识别工况类型。

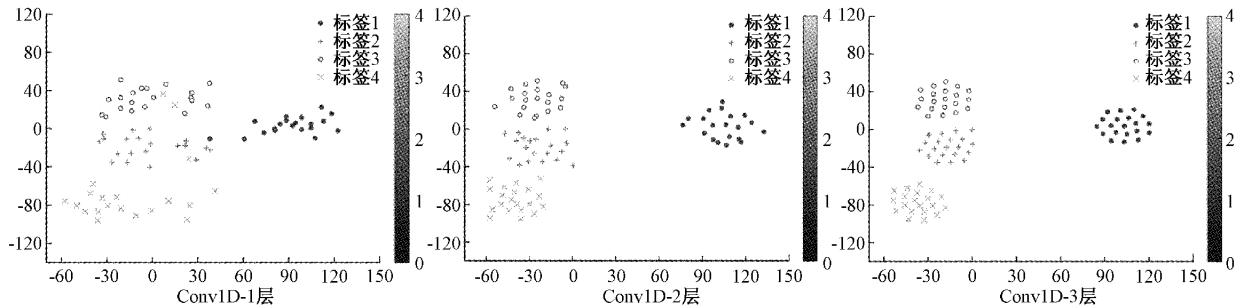


图7 t-分布散点图

3.2 与其他方法对比分析

管道泄漏识别常采用特征提取与构造分类器相结合完成分类任务,目前主流的分类方法主要有卷积神经网络和支持向量机等方法。为了验证EEMD-ICNN模型比常用的分类方法具有更高的识别准确率,将本文方法计算结果与EEMD-CNN及文献[22]模型进行对比,10次实验后3种模型的诊断效果如图8所示。

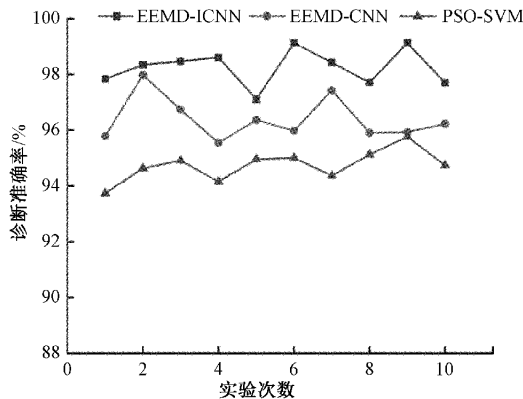


图8 3种模型10次诊断准确率对比

由图8可知,对于多工况下的管道泄漏识别问题,基于卷积神经网络模型的识别准确率均较高,能达到95%以上,尤其是ICNN模型在自适应学习和特征特取方面具有更为明显的优势,取得了高达98.25%的平均识别准确率,表明本文所提方法能够有效解决多工况下的泄漏识别问题。而支持向量机(support vector machine, SVM)分类效果仍有待提高,这是因为传统的浅层特征机器学习算法在处理多工况分类问题时特征提取能力有限,无法准确表征数据之间的映射关系,故识别准确率最低。

4 结 论

针对多工况下输油管道泄漏识别面临的原始数据预处理繁琐、误报率高等问题,提出了运用EEMD和ICNN

结合的分类方法。通过实验室模拟得到管道4种运行工况的压力信号,数据预处理后搭建识别模型进行泄漏诊断。主要结论如下:

基于EEMD的信号预处理方法能够有效地抑制信号模态混叠的影响,利用相关性分析筛选出可以充分表征管道运行工况的真实特征信息,为泄漏识别模型提供了良好的算法基础。

EEMD和ICNN组合模型在较少训练数据下依旧能取得良好的分类结果,提高了多工况下的管道泄漏识别准确率。分类结果显示,测试的80个样本中有1个样本结果与真实结果不符,模型的平均识别准确率为98.25%比SVM(94.74%)和未改进CNN(96.39%)模型有明显提升。

参考文献

- [1] 郎宪明,李平,曹江涛,等.长输油气管道泄漏检测与定位技术研究进展[J].控制工程,2018,25(4):621-629.
- [2] 郝永梅,杜璋昊,杨文斌,等.基于改进ELMD和多尺度熵的管道泄漏信号识别[J].中国安全科学学报,2019,29(8):105-111.
- [3] HAO Y M, DU Z, XUNG Z, et al. Urban hazardous chemicals pipeline leakage positioning method based on CELMD-MCKD [J]. Nondestructive Testing and Evaluation, 2021, 36(5):477-493.
- [4] 高文傲,刘洋,李登科,等.基于光纤监测技术的燃气管道泄漏研究[J].中国安全科学学报,2019,29(8):67-72.
- [5] NING F L, CHENG Z, MENG D, et al. A framework combining acoustic features extraction method and random forest algorithm for gas pipeline leak detection and classification[J]. Applied Acoustics, 2021, DOI: 10.1016/j.apacoust.2021.108255.
- [6] 高琳,曹建国.管道泄漏检测方法研究综述[J].现代制造工程,2022(2):154-162.

- [7] 李帅永,程振华,毛维培,等.基于改进经验小波变换及互谱相位差谱的供水管道泄漏声振动定位方法[J].仪器仪表学报,2019,40(12):80-91.
- [8] 文静,张敏姿,张恒.90°弯管存在下的供水管道泄漏定位研究[J].振动与冲击,2018,37(6):92-98.
- [9] 赵越,赵英亮,王黎明,等.基于 EEMD 的压力管道泄漏声源定位[J].国外电子测量技术,2021,40(7):128-132.
- [10] 王晓敏,骆正山,赵乐新.基于多层感知器神经网络和统计小波特征的管道泄漏诊断[J].安全与环境学报,2021,21(4):1483-1489.
- [11] 魏强,崔洪斌,谢耀国,等.海洋平台油气管道疲劳裂纹 AE 信号特征提取及识别研究[J].振动与冲击,2021,40(8):70-78.
- [12] XIE Y C, XIAO Y, LIU X, et al. Time-frequency distribution map-based convolutional neural network(CNN) model for underwater pipeline leakage detection using acoustic signals[J]. Sensors, 2020, 20(18): 5040-5040.
- [13] 唐秀家,颜大椿.基于神经网络的管道泄漏检测方法及其仪器[J].北京大学学报(自然科学版),1997(3):49-57.
- [14] 孙洁娣,王利轩,温江涛,等.改进稠密块轻量化神经网络的管道泄漏孔径识别[J].仪器仪表学报,2022,43(3):98-108.
- [15] BOHORQUEZ J, LAMBERT M F, ALEXANDER B, et al. Stochastic resonance enhancement for leak detection in pipelines using fluid transients and convolutional neural networks[J]. Journal of Water Resources Planning and Management, 2022, DOI: 10.1061/(ASCE)WR.1943-5452.0001504.
- [16] WU Q, LEE C M. A modified leakage localization method using multilayer perceptron neural networks in a pressurized gas pipe[J]. Applied Sciences, 2019, DOI:10.3390/app9091954.
- [17] 温江涛,付磊,孙洁娣,等.压缩感知结合卷积网络的天然气管道泄漏孔径识别[J].振动与冲击,2020,39(21):17-23.
- [18] 张瑞程,王新颖,胡磊磊,等.基于一维卷积神经网络的燃气管道泄漏声发射信号识别[J].中国安全生产科学技术,2021,17(2):104-109.
- [19] WU Z, HUANG N E. Ensemble empirical mode decomposition: A noise-assisted data analysis method[J]. Advances in Adaptive Data Analysis, 2009, 1(1): 1-41.
- [20] 梁伟,张来斌,王朝晖.基于数据挖掘的负压波特征参数优化方法[J].石油机械,2008,36(12):38-42,85-86.
- [21] 王望望,邓林峰,赵荣珍,等.集成 KPCA 与 t-SNE 的滚动轴承故障特征提取方法[J].振动工程学报,2021,34(2):431-440.
- [22] 王学渊,陈志刚,钟新荣,等.基于粒子群优化的 SVM 供水管道泄漏诊断方法[J].现代电子技术,2018,41(7):156-159,164.

作者简介

骆正山,教授,博士生导师,主要研究方向为油气管道风险评估与管理、机器学习、智能信息处理方向等。
E-mail:luozhengshan@163.com